USAARL Report No. 99-14

Notes on the Calculation of Power in Determining Sample Size and the Construction of Digram-Balanced Designs

By

Robert M. Wildzunas

Aircrew Health and Performance Division

March 1999

U.S. Army Aeromedical Research Laboratory
Fort Rucker, Alabama 36362-0577

Notice

Qualified requesters

Qualified requesters may obtain copies from the Defense Technical Information Center (DTIC), Cameron Station, Alexandria, Virginia 22314. Orders will be expedited if placed through the librarian or other person designated to request documents from DTIC.

Change of address

Organizations receiving reports from the U.S. Army Aeromedical Research Laboratory on automatic mailing lists should confirm correct address when corresponding about laboratory reports.

Disposition

Destroy this document when it is no longer needed. Do not return it to the originator.

Disclaimer

The views, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other official documentation. Citation of trade names in this report does not constitute an official Department of the Army endorsement or approval of the use of such commercial items.

Reviewed:

MORRIS R. LATTIMORE, JR.
Colonel, MS
Director, Aircrew Health &
Performance Division

JOHN A. CALDWELL, Ph.D.
Chairman, Scientific Review
Committee

Released for publication:

CHERRY L. GAFFNEY
Colonel, MC, SFS
Commanding

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release, distribution unlimited
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) USAARL Report No. 99-14	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION U.S. Army Aeromedical Research Laboratory 6b. OFFICE SYMBOL *(If applicable)* MCMR-UAS	7a. NAME OF MONITORING ORGANIZATION U.S. Army Medical Research and Materiel Command
6c. ADDRESS *(City, State, and ZIP Code)* P.O. Box 620577 Fort Rucker, AL 36362-0577	7b. ADDRESS *(City, State, and ZIP Code)* Fort Detrick Frederick, MD 21702-5012
8a. NAME OF FUNDING / SPONSORING ORGANIZATION 8b. OFFICE SYMBOL *(If applicable)*	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS *(City, State, and ZIP Code)*	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. | PROJECT NO. | TASK NO. | WORK UNIT ACCESSION NO.

11. TITLE *(Include Security Classification)*
Notes on the Calculation of Power in Determining Sample Size and the Construction of Digram-Balanced Designs (U)

12. PERSONAL AUTHOR(S)
Robert M. Wildzunas

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT *(Year, Month, Day)*	15. PAGE COUNT
Final	FROM 1999 TO	1999 March	24

16. SUPPLEMENTAL NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*
Experimental Design
Statistical power
Counterbalancing

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*
This paper presents step-by-step examples to help researchers determine the appropriate sample size for an experiment and to construct digram-balanced matrices for their designs. The power calculations are generic for fixed-effects analysis of variance (ANOVA) models and thus, are generalizable for the most common experimental designs employed at the U.S. Army Aeromedical Research Laboratory (USAARL). The mathematical algorithm for constructing the digram-balanced design works for both odd and even numbers of treatments; a USAARL-developed computer program using this algorithm is included for constructing and printing out the resulting counterbalanced matrix.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT X UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Chief, Science Support Center	22b. TELEPHONE *(Include Area Code)* (334) 255-6907 22c. OFFICE SYMBOL MCMR-UAX-SI

DD Form 1473, JUN 86 — *Previous editions are obsolete.*

SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

Table of contents

Introduction

Two frequently recurring experimental design questions concern power and counterbalancing. This paper presents step-by-step examples to help researchers determine the appropriate sample size and to construct digram-balanced matrices for their experimental designs. The power calculations are generic for fixed-effects one-way analysis of variance[1] (ANOVA) models and thus, are generalizable for the most common experimental designs employed at the U.S. Army Aeromedical Research Laboratory (USAARL). The mathematical algorithm for constructing digram-balanced designs is unique and works for both even and odd numbers of treatments (*a*). For more precise design-specific estimates of power, the interested researcher is directed to Keppel's (1991) *Design and analysis: A researchers handbook*, or Kraemer's (1987) *How many subjects?*, for in-depth treatments of these topics. A USAARL-developed computer program is included for constructing and printing out the digram-balanced matrices for any value of *a*.

Power calculation and the determination of sample size

Power quite simply is defined as 1-β and represents the probability of rejecting the null hypothesis given that it is false. The generally accepted benchmark for power in experimental research is 0.80 (i.e., an 80% chance of rejecting a false null hypothesis). Keppel (1991) states that "Low power is poor science – we waste time, energy, and resources whenever we conduct an experiment that has a low probability of producing a significant result." Thus, the modern researcher often is asked to "estimate" the power of an experiment during early protocol development. This seemingly daunting task can be accomplished with a few relatively simple calculations.

Before we can calculate power it is necessary to determine Omega-squared (ω^2). Omega-squared is one of the most popular indices of effect size, other indices such as the squared multiple correlation coefficient (R^2) also are used, but for simplicity are not discussed here[2]. Omega-squared can be estimated in several ways. Cohen (1977) suggests the following rules-of-thumb for the behavioral and social sciences.

A "small" effect is an experiment that produces a ω^2 of 0.01 to 0.05
A "meduim" effect is an experiment that produces a ω^2 of 0.06 to 0.14.
A "large" effect is an experiment that produces a ω^2 of 0.15 or greater.

The values above generally are too conservative and are based on the assumption that the error variance found in most behavioral experiments is relatively large. Consequently, the number of subjects required for any given effect size can be quite large to obtain sufficient power (table 1).

---

1 For more complex designs, the smallest univariate F-ratio can be used conservatively to estimate power.

2 The difference between ω^2 and R^2 is generally small, and decreases as sample size (n) increases. Keppel (1991) directs interested readers to Maxwell, Camp, and Arvey (1981) for a discussion of these two indices.

Table 1. Sample size (n) as a function of power and effect size, for α=.05 (Keppel, 1991).

Effect Size (ω^2)	Power (1-β)								
	.10	.20	.30	.40	.50	.60	.70	.80	.90
0.01	21	53	83	113	144	179	219	271	354
0.06	5	10	14	19	24	30	36	44	57
0.15	3	5	6	8	10	12	14	17	22

For more precise values, the researcher can calculate ω^2 rather easily using the formula below:

$$\omega^2 = \frac{(a-1)(F-1)}{(a-1)(F-1)+(a)(n)}$$

Where: *a* is the number of treatments, *F* is the observed or estimated F-ratio value from a previous, similar experiment on the treatment of interest, and *n* is the number of subjects per treatment.

Once ω^2 has been calculated, ϕ is calculated to determine power from the Pearson-Hartley power planning tables (appendix A).

$$\phi = \sqrt{n \frac{\omega^2}{1-\omega^2}}$$

EXAMPLE 1: Suppose we wanted to know if 8 subjects[3] would yield sufficient power at the α=.05 confidence level in an experiment with 4 treatments (or 4 levels of one treatment). The observed (or estimated) F-ratio based on previous, similar experiments with these treatments is 5.75.

$$\omega^2 = \frac{(a-1)(F-1)}{(a-1)(F-1)+(a)(n)} = \frac{(4-1)(5.75-1)}{(4-1)(5.75-1)+(4)(8)} = \frac{(3)(4.75)}{(3)(4.75)+(32)}$$

$$= \frac{(14.25)}{(14.25)+(32)} = \frac{(14.25)}{(46.25)} \cong .3081$$

$$\phi = \sqrt{n\frac{\omega^2}{1-\omega^2}} = \sqrt{8\frac{.3081}{1-.3081}} = \sqrt{8\frac{.3081}{.6919}} = \sqrt{8(.4453)} = \sqrt{3.56} \cong 1.88$$

[3] For the purpose of these calculations, it does not matter whether the design is within- or between-subjects as *n* represents the number of subjects per treatment. Thus, 32 subjects, 8 different subjects in each treatment group, is equivalent to eight subjects, each tested four times, once within each treatment group. However, this design issue should be considered when estimating *F*; within-subject designs generally result in reduced error variance which in turn results in a larger F-ratio.

These values then are looked up in the appropriate Pearson-Hartley (1972) chart (appendix A). The various charts are distinguished by the degrees of freedom in the numerator of the F-ratio ($df_{num} = a$-1). Each chart contains two sets of power curves for 11 different degrees of freedom in the denominator of the F-ratio [$df_{denom}=(a)(n\text{-}1)$]. Values not found in the chart may be interpolated between the two closest values. On most of the charts, the first set of curves represents values of α=.05 and the second set of curves represents α=.01[3].

Thus, given the previous example, we would consult the df_{num}=3 chart (a=4, hence a-1=3) (see figure). Then, using the first set of curves (α=0.05), we would find the calculated value of ϕ ($\cong$1.88) on the abscissa (x-axis) and trace upward to intersect the df_{denom}=28 [$(a)(n\text{-}1)$, or 4(7)=28, interpolated just below the df_{denom}=30 line]. Finally, tracing across to the ordinate (y-axis), the scale shows the resulting power value to be approximately 0.84. Since we were shooting for a power of 0.80, we can conclude that 8 subjects per treatment should yield sufficient power for the design in question. However, had the values been too low, we would have had to recycle through the calculations until an adequate sample size was determined.

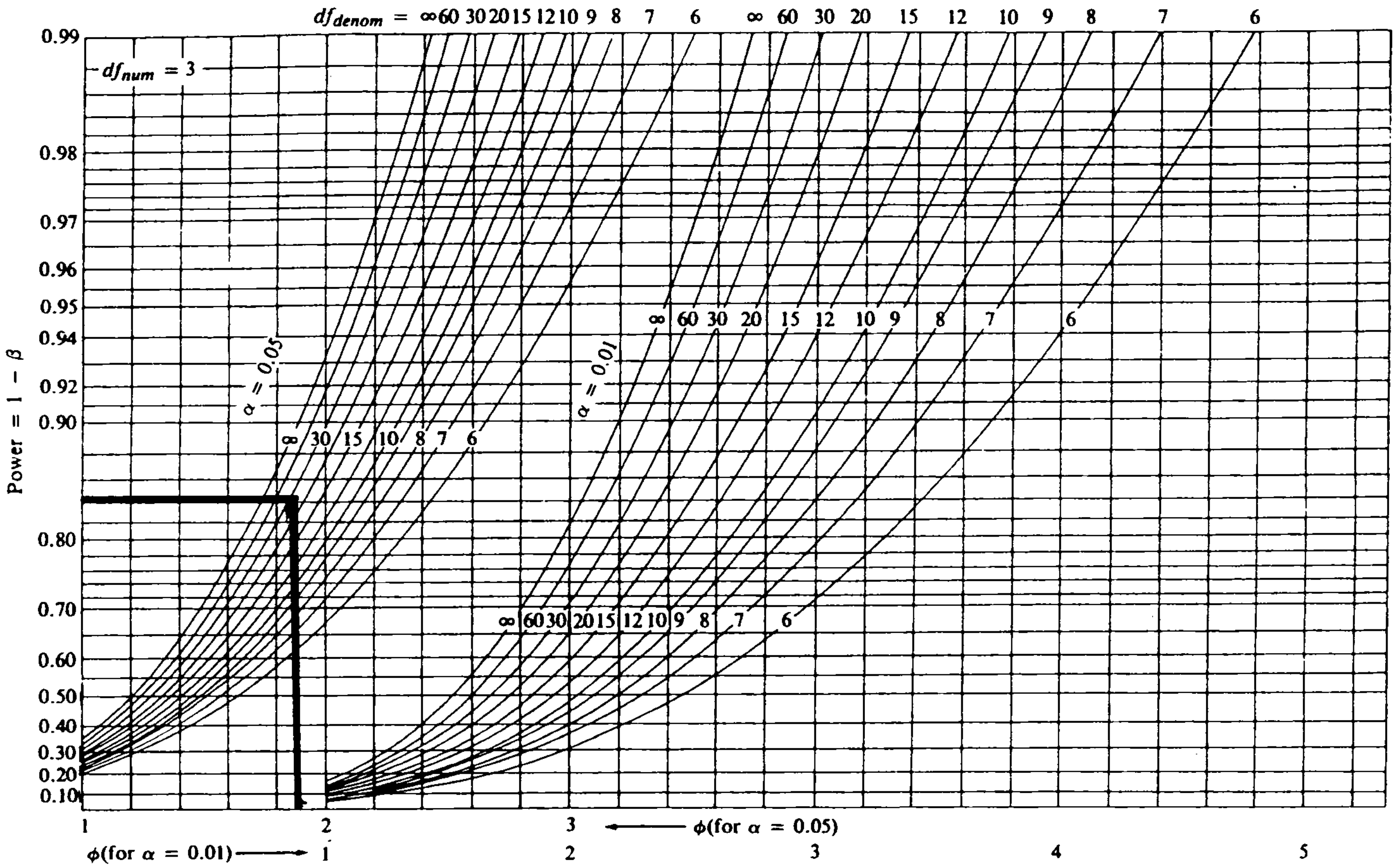


Figure. Pearson-Hartley power planning chart for df_{num}=3

[3] Please note that this generalization is reversed only for the first chart (df_{num}=1).

Constructing digram-balanced designs

Many research designs call for the subject to perform a task repeatedly. In such cases, the investigator systematically manipulates one or more of the independent variables to determine the treatment effects on the subject. Notwithstanding, investigators must be careful not to confuse any treatment effects with order effects (i.e., practice or fatigue effects). Practice effects exist when, through repeated exposure to the test condition, subjects become familiar with the experiment. They may relax and do a little better, they may get better using the apparatus, or they may develop strategies for dealing with problems presented. Any of these may have a positive effect on performance that is not caused by changes in the independent variable. Likewise, with fatigue effects, subjects can become bored, irritated, or tired, all of which can have a negative effect on performance.

Counterbalancing is a technique to control for order effects that distributes this source of error across the treatment conditions. Many methods for counterbalancing have been developed and some are more desirable than others[5]. Probably the least desirable is the cyclic method in which the order of the conditions is varied but the sequence is the same regardless of the number of treatments "a" (table 2). Although this type of counterbalancing satisfies the requirement that each treatment appear only once in any position, the primary problem (in this example) is that 4 always follows 3, which always follows 2, which always follows 1, and so on. Thus, if one treatment has a carryover effect on another treatment that follows it, the experiment is still confounded.

Table 2. Sequence counterbalancing.

Subject	Testing sequence			
1	1	2	3	4
2	4	1	2	3
3	3	4	1	2
4	2	3	4	1

Classical Latin squares on the other hand are a little better for dealing with these confounds except that you will notice that some parts of the sequence tend to repeat themselves (table 3). Digram-balanced Latin squares (Wagenaar, 1969) offer a better arrangement in that each treatment occurs in each position only once, and that each treatment is preceded or followed by each other treatment only once (table 4). Wagenaar presents a simple method for constructing digram-balanced Latin squares that superficially seems to meet all the requirements for counterbalancing. However, he notes that these solutions exist only for even values of a. He goes on to state that, "a general mathematical theory which gives a simple rule for a given value of N [sic, for this discussion N=*a*] and which accounts for the non-existence of solutions for odd values of N has not been found in spite of many efforts."

[5] All examples are for a four treatment experiment such as the one outlined in the power discussion.

Table 3. Classic Latin square.

Subject	Testing Sequence			
1	1	2	3	4
2	2	1	4	3
3	3	4	1	2
4	4	3	2	1

Table 4. Digram-balanced Latin square.

Subject	Testing sequence			
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

Namboodiri (1972) solved the odd treatment dilemma using an intricate "extra period" design that started with a cyclic Latin square in which each row was interlaced with its own mirror image resulting in an $n \times 2n$ arrangement. For example, for n=4, the first row would be 1, 4, 2, 3, 3, 2, 4, 1. When sliced down the middle, each half represented a period, each row the subjects, and each number within the square the treatment. A simpler solution, also utilizing an extra period (only for odd numbers) is presented here.

This digram-balance method begins with an interlaced row of numbers where $n=a$. The sequence of these numbers for all levels of a follows the algorithm:

> 1, 2, n, 3, n-1, 4, n-2, ... until all values of 1 to n are exhausted.

Thus, for the following levels of a (1 to 10) the first row would be:

a=1 1
a=2 1, 2
a=3 1, 2, 3
a=4 1, 2, 4, 3
a=5 1, 2, 5, 3, 4
a=6 1, 2, 6, 3, 5, 4
a=7 1, 2, 7, 3, 6, 4, 5
a=8 1, 2, 8, 3, 7, 4, 6, 5
a=9 1, 2, 9, 3, 8, 4, 7, 5, 6
and,
a=10 1, 2, 10, 3, 9, 4, 8, 5, 7, 6 and so on . . .

Then, each column is filled sequentially from the number in the first row to n. Following n, the sequence begins again with 1, 2, ... until all values from 1 to n are exhausted. If the number of treatments, a, is even, there are no additional steps (table 5). However, if a is odd, each row is repeated in reverse order to create the extra period required to ensure that each treatment is preceeded by each other treatment the same number of times (table 6).

Table 5. Digram-balanced a=4 (even).

Subject	Testing sequence			
1	1	2	4	3
2	2	3	1	4
3	3	4	2	1
4	4	1	3	2

Step 1. Use algorithm to complete first row.

Step 2. Fill in columns sequentially.

Table 6. Digram-balanced a=5 (odd).

Subject	Testing sequence				
1	1	2	5	3	4
2	2	3	1	4	5
3	3	4	2	5	1
4	4	5	3	1	2
5	5	1	4	2	3
6	4	3	5	2	1
7	5	4	1	3	2
8	1	5	2	4	3
9	2	1	3	5	4
10	3	2	4	1	5

First period
(same procedure as above)

Second period
(rows reverse-ordered from first period)

Appendix B contains digram-balanced squares for 1 to 10 treatments. The squares obtained represent the minimum arrangement to counterbalance the design. The rows represent the number of subjects required. Note that for odd numbered treatments n=2a. Thus, to ensure digram-balancing for even numbers of treatments, n required is a multiple of a. Whereas, for odd numbers of treatments, n required is a multiple of 2a. These facts can affect your final design.

EXAMPLE 2: Suppose from your power calculations for a=5 treatments, you determined that 12 subjects are required to yield sufficient power (power = .86). Since you suspect strong order effects, you realize that in order to counterbalance the design you need at least 20 subjects (and these subjects are expensive in both time and resources). After careful consideration, you decided it may be possible to eliminate one level of treatment a and run only 12 subjects (three identical periods of a=4). But first, you recycled through the calculations for 4 treatments and determined that power (.82), although lower, was still acceptable.

The accompanying diskette will produce digram-balanced matrices for any number of treatments (appendix C).

References

Cohen, J. 1977. Statistical power analysis for the behavioral sciences (rev. ed.). New York, NY: Academic Press.

Keppel, G. 1991. Design and analysis: a researcher's handbook. Englewood Cliffs, NJ: Prentice Hall.

Kraemer, H.C. 1987. How Many Subjects? Astatistical power analysis in research. Newbury Park, U.K.: Sage Publications.

Maxwell, S.E., Camp, C.J., and Arvey, R.D. 1981. Measures of strength of associations: A comparative examination. Journal of Applied Psychology, 66: 525-534.

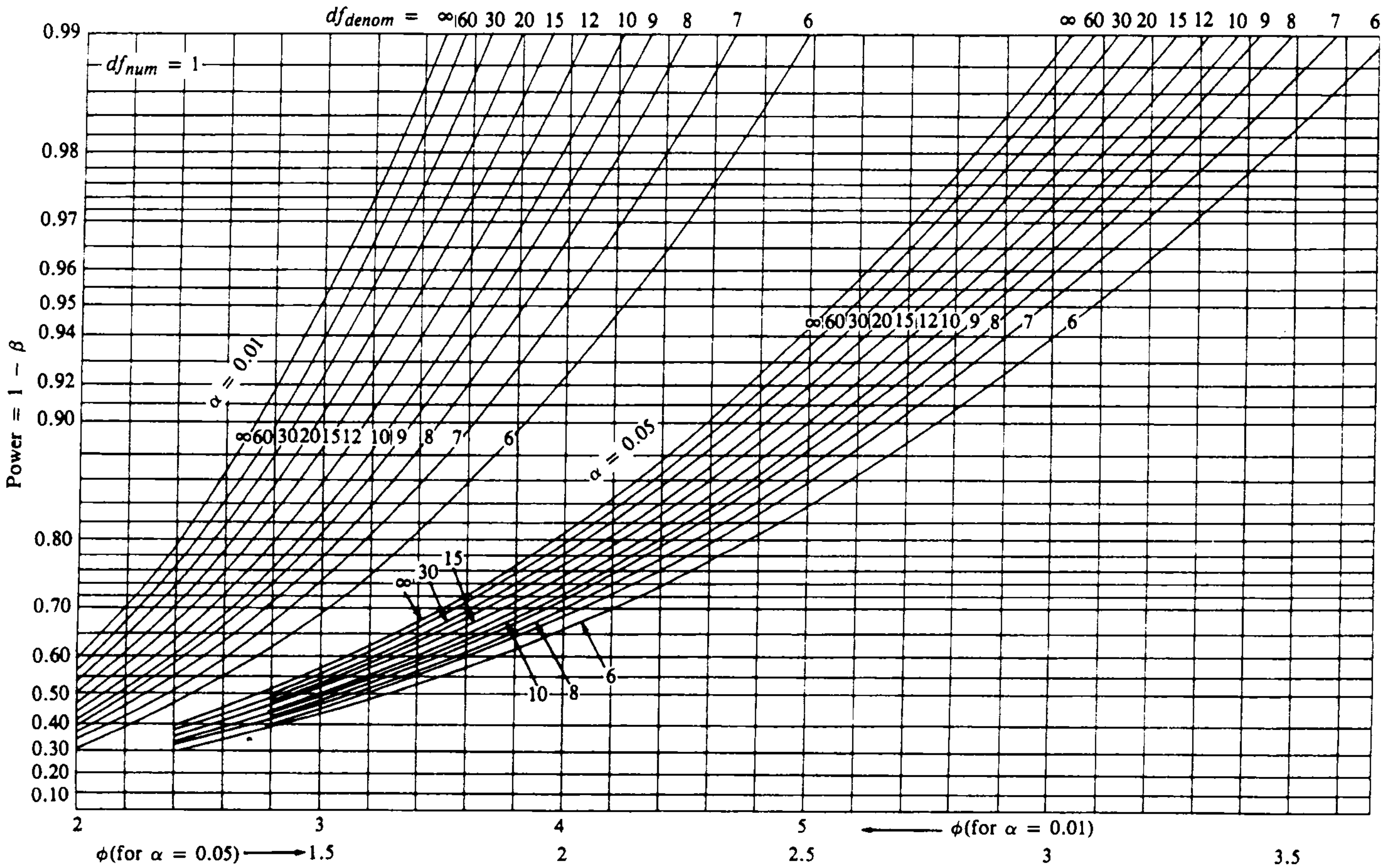
Pearson, E.S., and Hartley, H.O. (Eds). 1972. Biometrika tables for statisticians (3rd ed. , Vol. 1). New York, NY: Cambridge University Press.

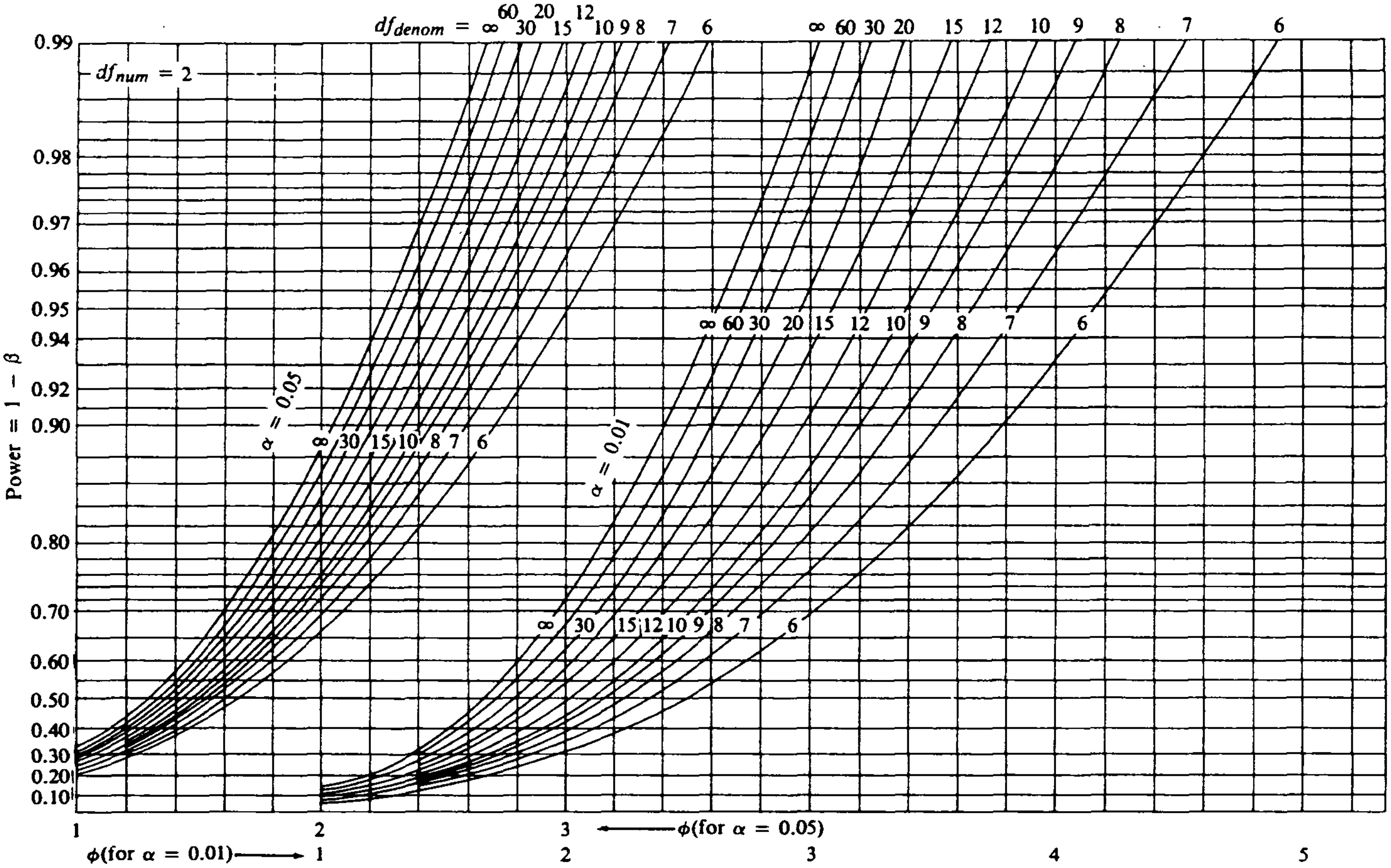
Namboodiri, N.K. 1972. Experimental designs in which each subject is used repeatedly. Psychological Bulletin, 77: 54-64.

Wagenaar, W.A. 1969. Note on the construction of digram-balanced Latin squares. Psychological Bulletin, 72: 384-386.

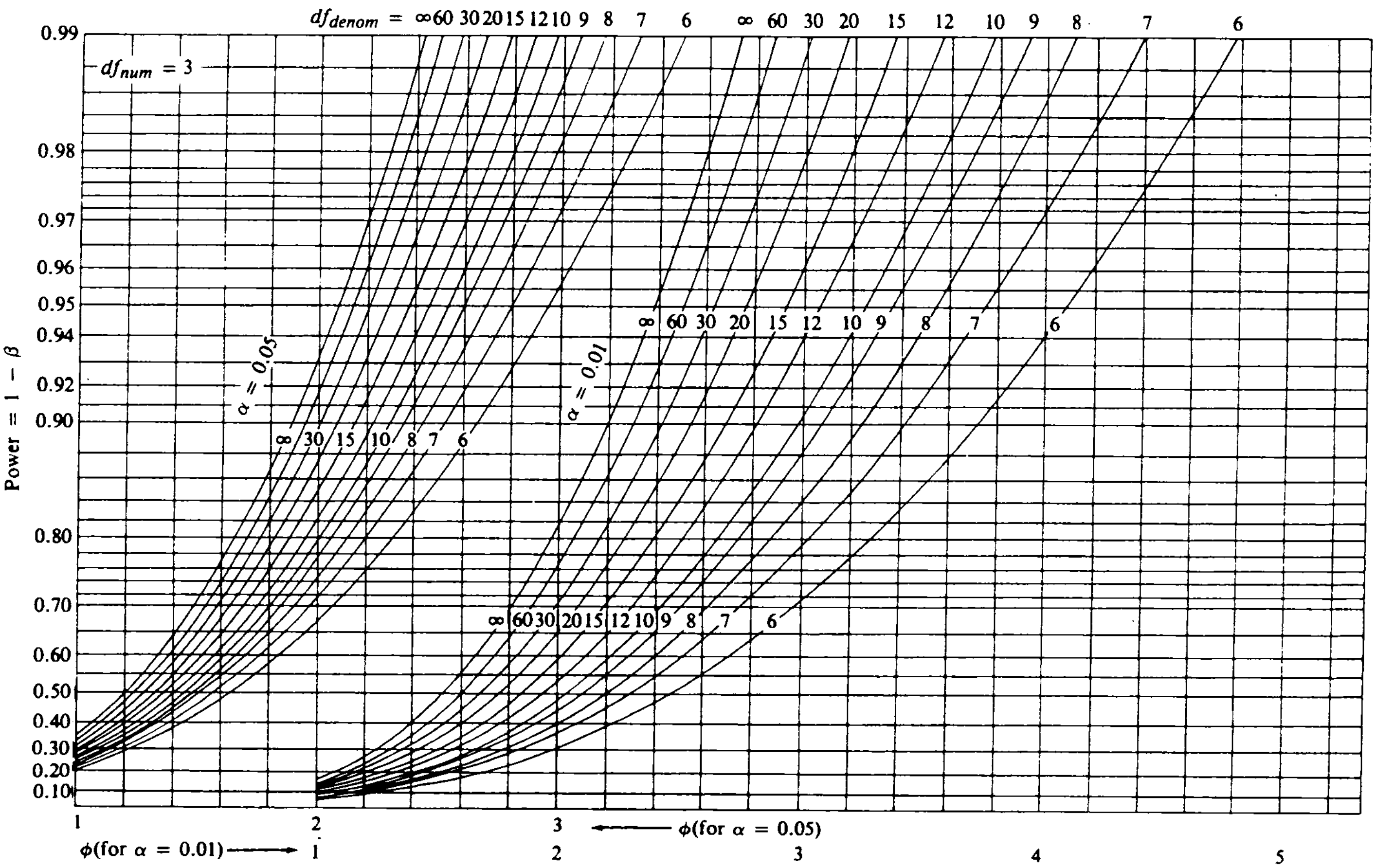
Appendix A.

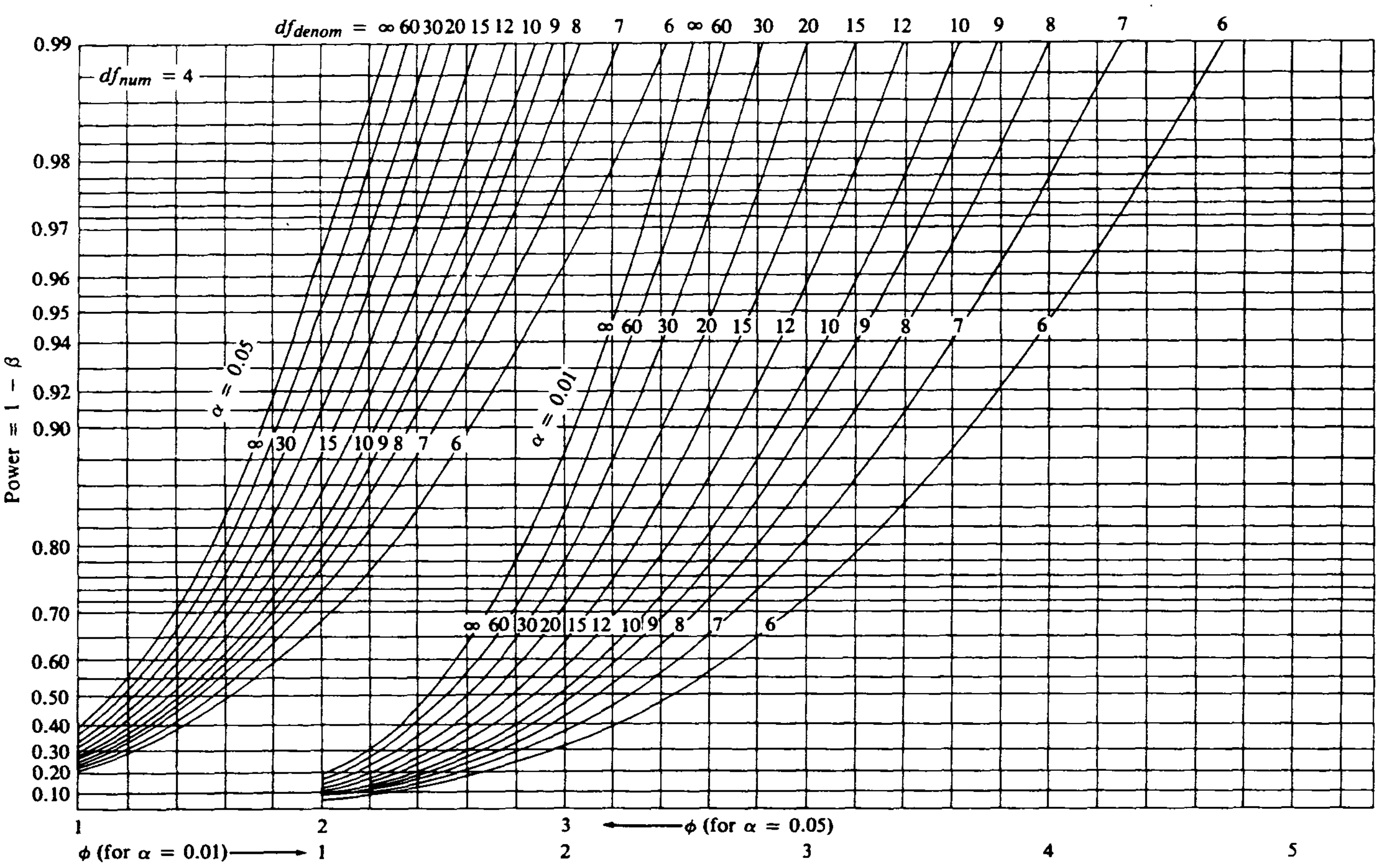
Pearson-Hartley power planning charts (Pearson and Hartley, 1972).



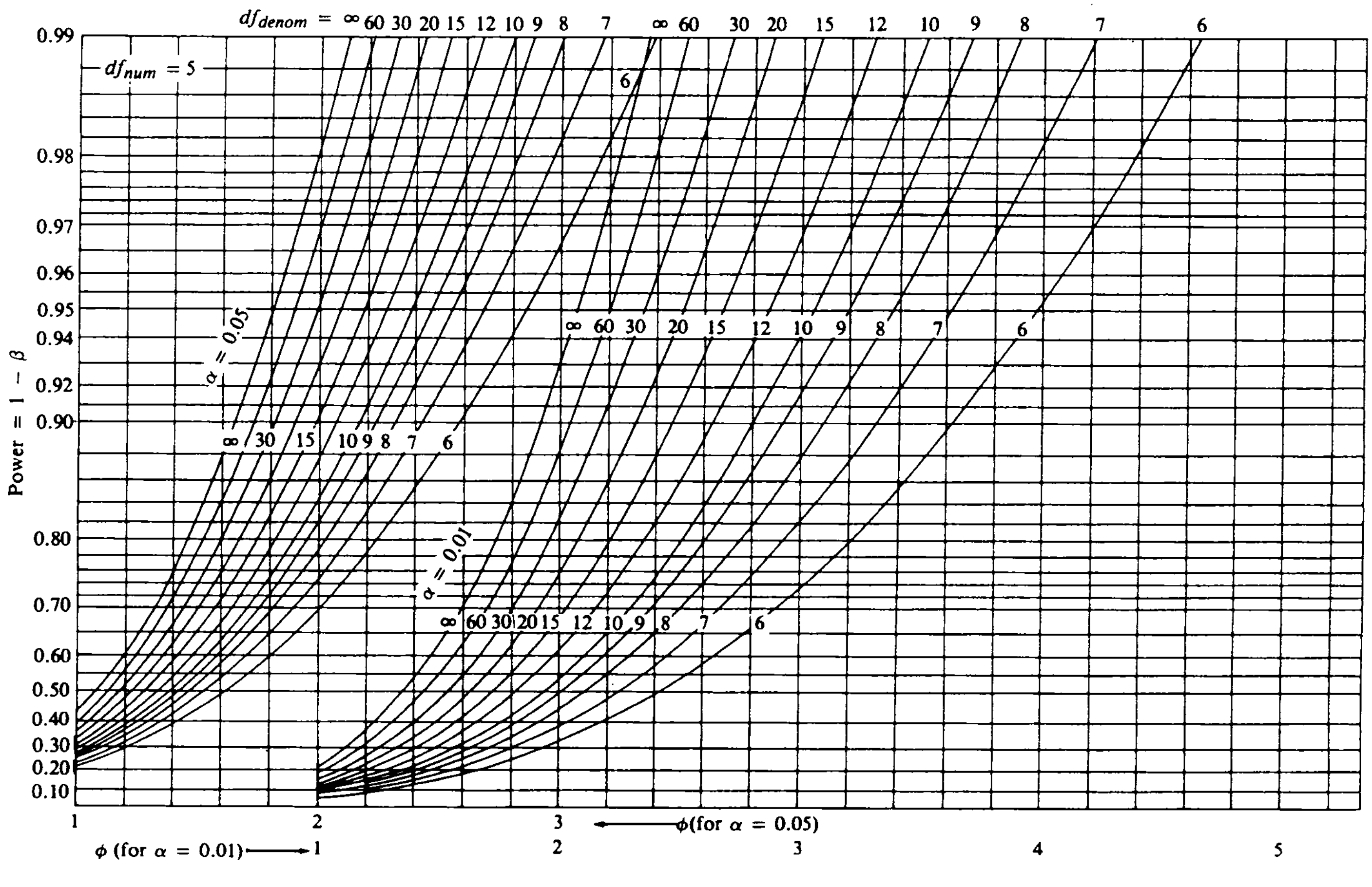


df_{denom} = ∞ 60 30 20 15 12 10 9 8 7 6
∞ 60 30 20 15 12 10 9 8 7 6
$df_{num} = 2$
Power = 1 − β
α = 0.05
α = 0.01
∞ 30 15 10 8 7 6
∞ 60 30 20 15 12 10 9 8 7 6
∞ 30 15 12 10 9 8 7 6
0.99
0.98
0.97
0.96
0.95
0.94
0.92
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
1 2 3 ← φ(for α = 0.05)
φ(for α = 0.01) → 1 2 3 4 5

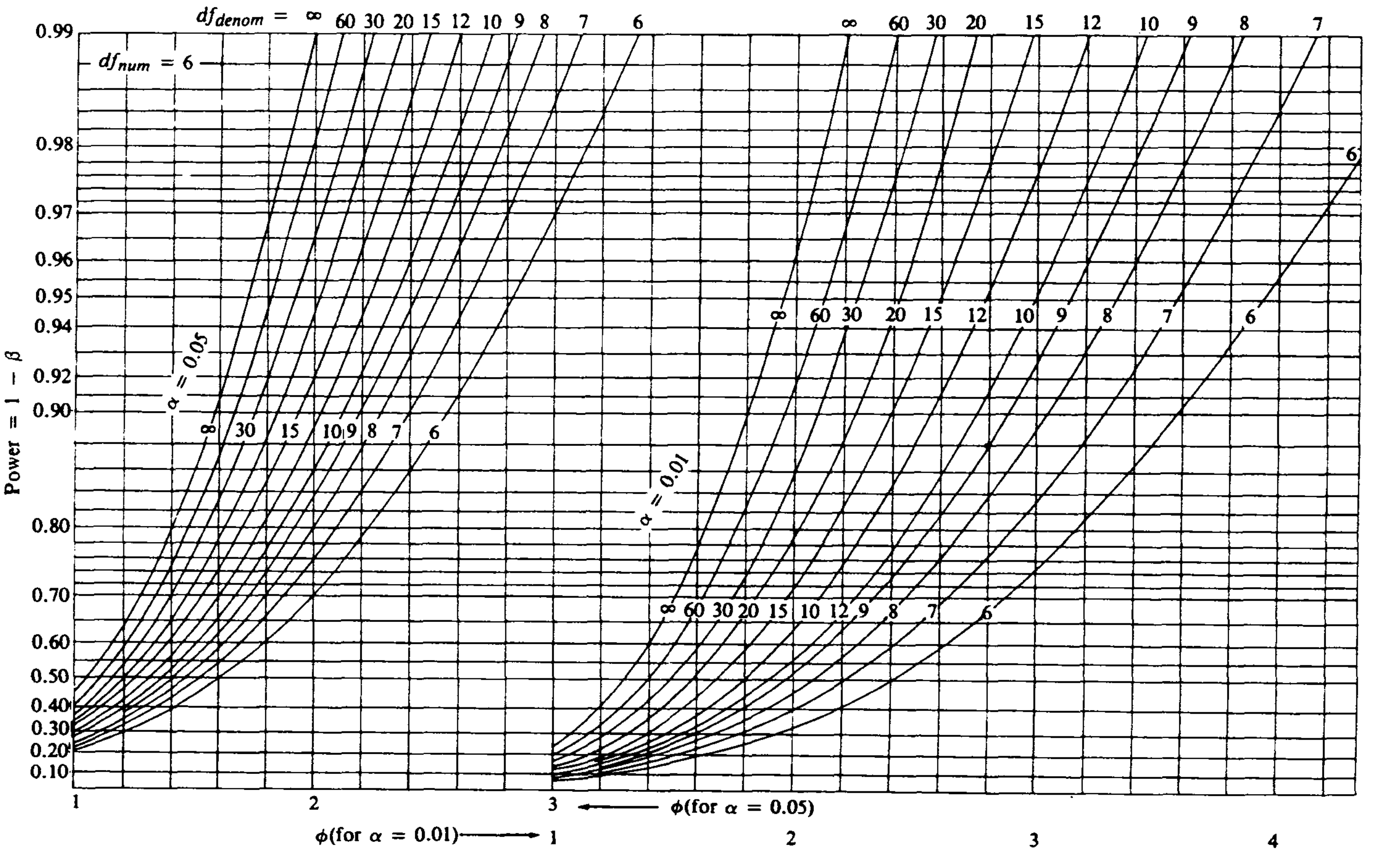

df_{denom} = ∞ 60 30 20 15 12 10 9 8 7 6 ∞ 60 30 20 15 12 10 9 8 7 6
df_{num} = 3
Power = 1 − β
α = 0.05
∞ 30 15 10 8 7 6
α = 0.01
∞ 60 30 20 15 12 10 9 8 7 6
∞ 60 30 20 15 12 10 9 8 7 6
0.99 0.98 0.97 0.96 0.95 0.94 0.92 0.90 0.80 0.70 0.60 0.50 0.40 0.30 0.20 0.10
1 2 3 ← φ(for α = 0.05)
φ(for α = 0.01) → 1 2 3 4 5

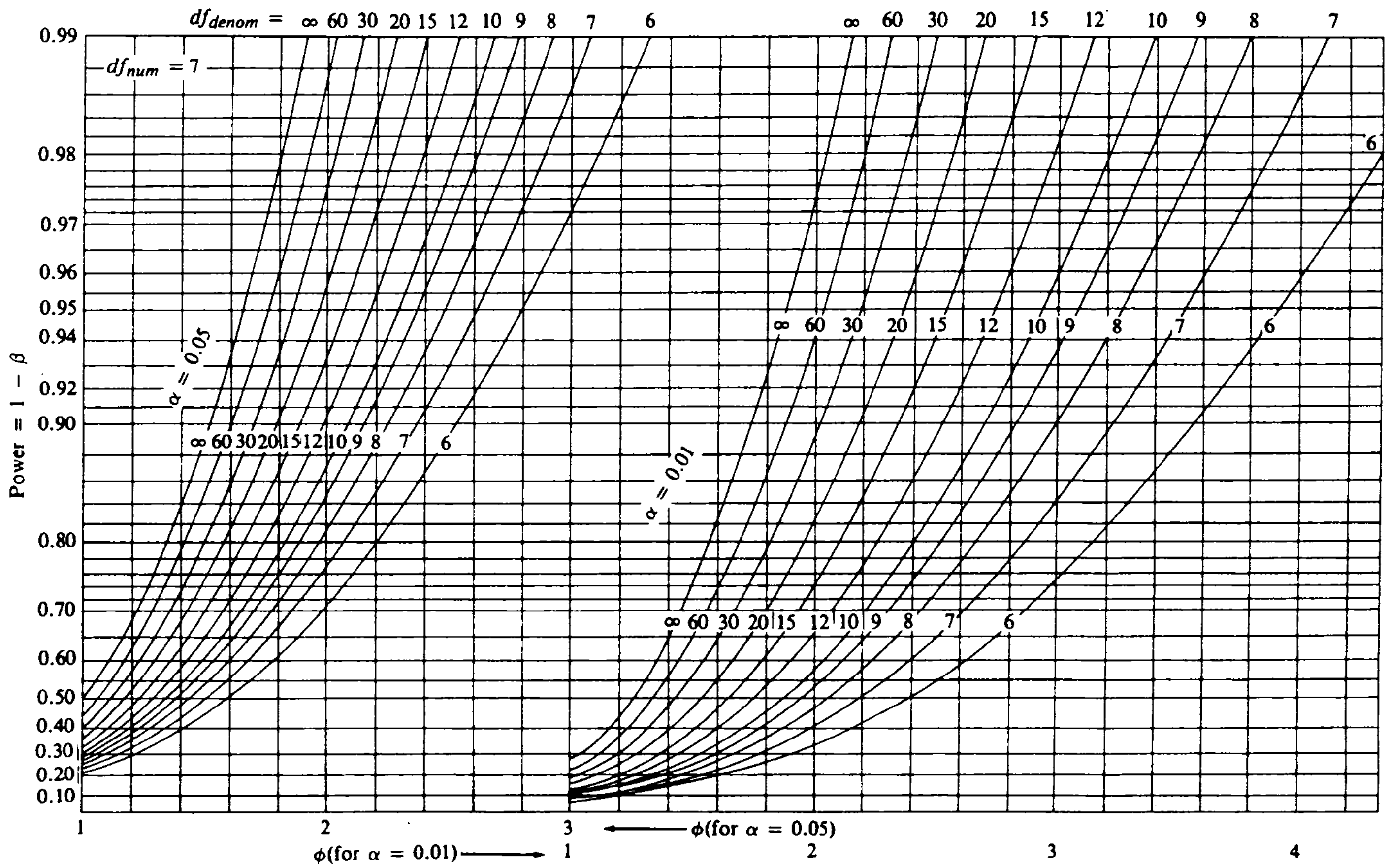


df_{denom} = ∞ 60 30 20 15 12 10 9 8 7 6 ∞ 60 30 20 15 12 10 9 8 7 6
df_{num} = 4
Power = 1 − β
0.99
0.98
0.97
0.96
0.95
0.94
0.92
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
α = 0.05
∞ 30 15 10 9 8 7 6
α = 0.01
∞ 60 30 20 15 12 10 9 8 7 6
∞ 60 30 20 15 12 10 9 8 7 6
1 2 3 ← φ (for α = 0.05)
φ (for α = 0.01) → 1 2 3 4 5

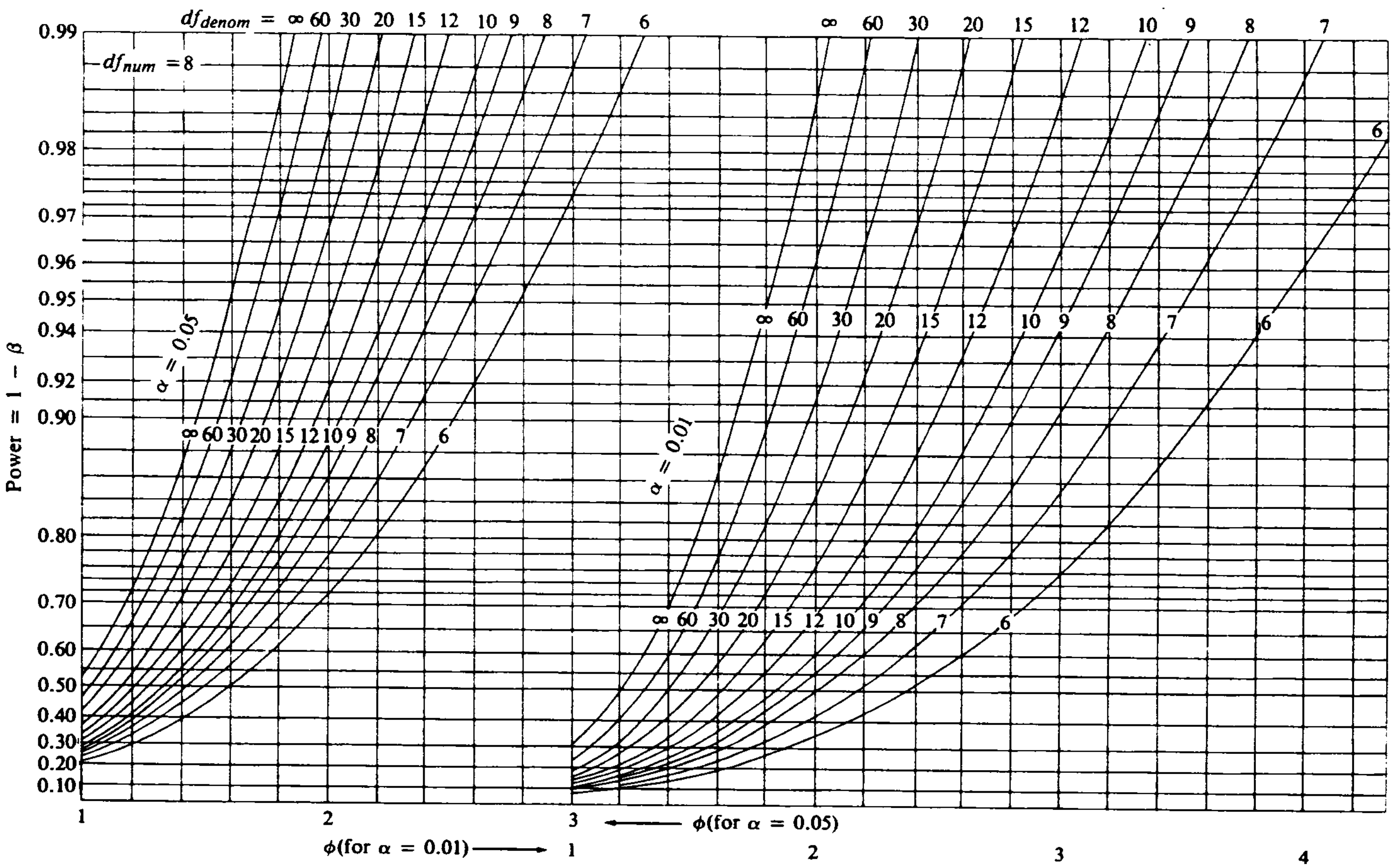


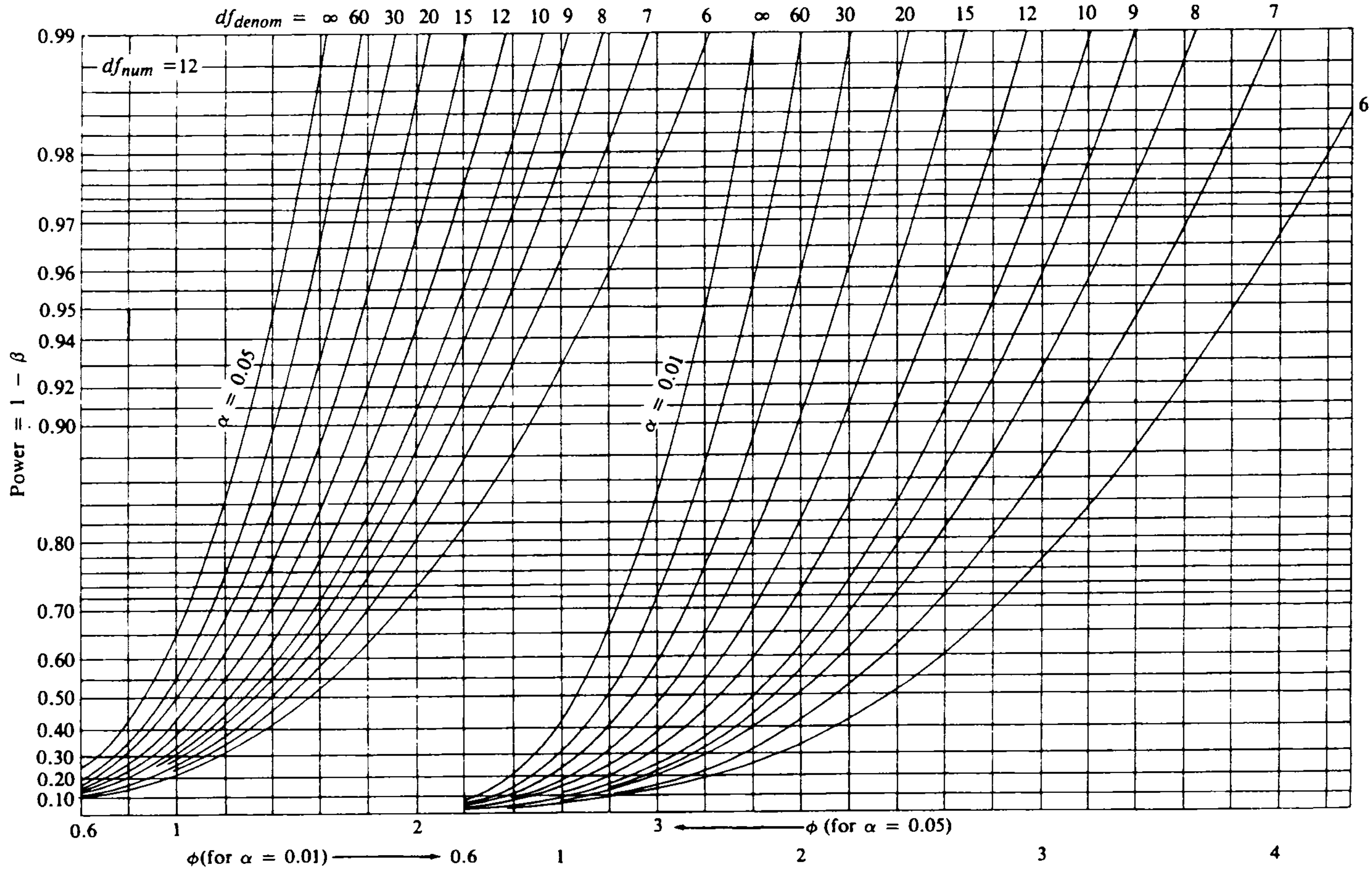
df_{denom} = ∞ 60 30 20 15 12 10 9 8 7 ∞ 60 30 20 15 12 10 9 8 7 6
$df_{num} = 5$
$\alpha = 0.05$
$\alpha = 0.01$
∞ 30 15 10 9 8 7 6
∞ 60 30 20 15 12 10 9 8 7 6
∞ 60 30 20 15 12 10 9 8 7 6
Power = $1 - \beta$
0.99
0.98
0.97
0.96
0.95
0.94
0.92
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
ϕ (for α = 0.05)
ϕ (for α = 0.01)


df_{denom} = ∞ 60 30 20 15 12 10 9 8 7 6
∞ 60 30 20 15 12 10 9 8 7
df_{num} = 6
0.99
0.98
0.97
0.96
0.95
0.94
0.92
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
Power = 1 − β
α = 0.05
∞ 30 15 10 9 8 7 6
α = 0.01
∞ 60 30 20 15 12 10 9 8 7 6
∞ 60 30 20 15 10 12 9 8 7 6
1
2
3
φ(for α = 0.05)
φ(for α = 0.01)
1
2
3
4

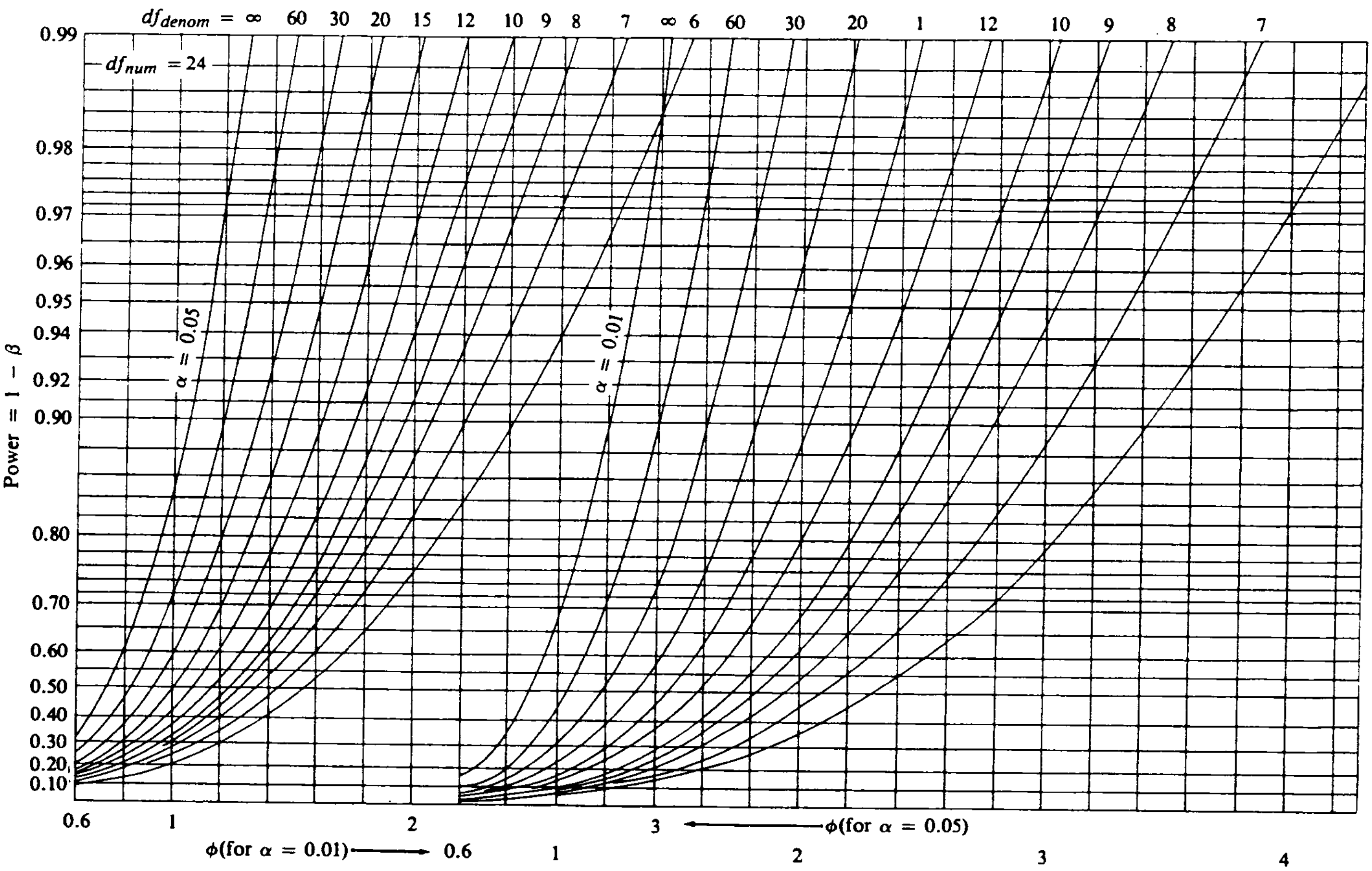


df_{denom} = ∞ 60 30 20 15 12 10 9 8 7 6
∞ 60 30 20 15 12 10 9 8 7
df_{num} = 7
$\alpha = 0.05$
$\alpha = 0.01$
Power = 1 − β
0.99
0.98
0.97
0.96
0.95
0.94
0.92
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
ϕ(for α = 0.01)
ϕ(for α = 0.05)


df_{denom} = ∞ 60 30 20 15 12 10 9 8 7 6
df_{num} = 8
Power = 1 − β
α = 0.05
α = 0.01
0.99
0.98
0.97
0.96
0.95
0.94
0.92
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
φ(for α = 0.01)
φ(for α = 0.05)
1
2
3
4



df_{denom} = ∞ 60 30 20 15 12 10 9 8 7 6 ∞ 60 30 20 15 12 10 9 8 7
6
df_{num} = 12
α = 0.05
α = 0.01
Power = 1 − β
0.99
0.98
0.97
0.96
0.95
0.94
0.92
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.6 1 2 3 ← φ (for α = 0.05)
φ(for α = 0.01) → 0.6 1 2 3 4


df_{denom} = ∞ 60 30 20 15 12 10 9 8 7 ∞ 6 60 30 20 1 12 10 9 8 7
df_{num} = 24
α = 0.05
α = 0.01
Power = 1 − β
0.99
0.98
0.97
0.96
0.95
0.94
0.92
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.6 1 2 3 ← φ(for α = 0.05)
φ(for α = 0.01) → 0.6 1 2 3 4

Appendix B.

Digram-balanced designs for 1 to 10 treatments.

a=1	
Subject	Testing sequence
1	1

a=2		
Subject	Testing sequence	
1	1	2
2	2	1

a=3			
Subject	Testing sequence		
1	1	2	3
2	2	3	1
3	3	1	2
4	3	2	1
5	1	3	2
6	2	1	3

a=4				
Subject	Testing sequence			
1	1	2	4	3
2	2	3	1	4
3	3	4	2	1
4	4	1	3	2

a=5					
Subject	Testing sequence				
1	1	2	5	3	4
2	2	3	1	4	5
3	3	4	2	5	1
4	4	5	3	1	2
5	5	1	4	2	3
6	4	3	5	2	1
7	5	4	1	3	2
8	1	5	2	4	3
9	2	1	3	5	4
10	3	2	4	1	5

a=6						
Subject	Testing sequence					
1	1	2	6	3	5	4
2	2	3	1	4	6	5
3	3	4	2	5	1	6
4	4	5	3	6	2	1
5	5	6	4	1	3	2
6	6	1	5	2	4	3

a=7							
Subject	Testing sequence						
1	1	2	7	3	6	4	5
2	2	3	1	4	7	5	6
3	3	4	2	5	1	6	7
4	4	5	3	6	2	7	1
5	5	6	4	7	3	1	2
6	6	7	5	1	4	2	3
7	7	1	6	2	5	3	4
8	5	4	6	3	7	2	1
9	6	5	7	4	1	3	2
10	7	6	1	5	2	4	3
11	1	7	2	6	3	5	4
12	2	1	3	7	4	6	5
13	3	2	4	1	5	7	6
14	4	3	5	2	6	1	7

a=8								
Subject	Testing sequence							
1	1	2	8	3	7	4	6	5
2	2	3	1	4	8	5	7	6
3	3	4	2	5	1	6	8	7
4	4	5	3	6	2	7	1	8
5	5	6	4	7	3	8	2	1
6	6	7	5	8	4	1	3	2
7	7	8	6	1	5	2	4	3
8	8	1	7	2	6	3	5	4

a=9									
Subject	Testing sequence								
1	1	2	9	3	8	4	7	5	6
2	2	3	1	4	9	5	8	6	7
3	3	4	2	5	1	6	9	7	8
4	4	5	3	6	2	7	1	8	9
5	5	6	4	7	3	8	2	9	1
6	6	7	5	8	4	9	3	1	2
7	7	8	6	9	5	1	4	2	3
8	8	9	7	1	6	2	5	3	4
9	9	1	8	2	7	3	6	4	5
10	6	5	7	4	8	3	9	2	1
11	7	6	8	5	9	4	1	3	2
12	8	7	9	6	1	5	2	4	3
13	9	8	1	7	2	6	3	5	4
14	1	9	2	8	3	7	4	6	5
15	2	1	3	9	4	8	5	7	6
16	3	2	4	1	5	9	6	8	7
17	4	3	5	2	6	1	7	9	8
18	5	4	6	3	7	2	8	1	9

a=10										
Subject	Testing sequence									
1	1	2	10	3	9	4	8	5	7	6
2	2	3	1	4	10	5	9	6	8	7
3	3	4	2	5	1	6	10	7	9	8
4	4	5	3	6	2	7	1	8	10	9
5	5	6	4	7	3	8	2	9	1	10
6	6	7	5	8	4	9	3	10	2	1
7	7	8	6	9	5	10	4	1	3	2
8	8	9	7	10	6	1	5	2	4	3
9	9	10	8	1	7	2	6	3	5	4
10	10	1	9	2	8	3	7	4	6	5

Appendix C.

Balance diskette.